Limitations to Fréchet's Metric Embedding Method

Yair Bartal[*]   Nathan Linial[†]   Manor Mendel[‡]   Assaf Naor

Abstract

Fréchet's classical isometric embedding argument has evolved to become a major tool in the study of metric spaces. An important example of a Fréchet embedding is Bourgain's embedding [4]. The authors have recently shown [2] that for every $\varepsilon > 0$ any n-point metric space contains a subset of size at least $n^{1-\varepsilon}$ which embeds into ℓ_2 with distortion $O\left(\frac{\log(2/\varepsilon)}{\varepsilon}\right)$. The embedding used in [2] is non-Fréchet, and the purpose of this note is to show that this is not coincidental. Specifically, for every $\epsilon > 0$, we construct arbitrarily large n-point metric spaces, such that the distortion of any Fréchet embedding into ℓ_p on subsets of size at least $n^{1/2+\epsilon}$ is $\Omega\left((\log n)^{1/p}\right)$.

1 Introduction

Given two metric spaces (X, d_X), (Y, d_Y) and an embedding $f : X \to Y$, the *distortion* of f is defined as:
$$\mathrm{dist}(f) = \sup_{\substack{x,y \in X \\ x \neq y}} \frac{d_Y(f(x), f(y))}{d_X(x,y)} \cdot \sup_{\substack{x,y \in X \\ x \neq y}} \frac{d_X(x,y)}{d_Y(f(x), f(y))}.$$

We denote by $c_Y(X)$ the least distortion with which X may be embedded in Y. When $c_Y(X) \leq \alpha$ we say that X α-embeds into Y and denote $X \overset{\alpha}{\hookrightarrow} Y$. When there is a bijection f between two metric spaces X and Y with $\mathrm{dist}(f) \leq \alpha$ we say that X and Y are α-equivalent. For a class of metric spaces $\mathcal{M}$, $c_{\mathcal{M}}(X)$ is the minimum α such that X α-embeds into some metric space in $\mathcal{M}$. For $p \geq 1$ we denote $c_{\ell_p}(X)$ by $c_p(X)$. The parameter $c_2(X)$ is known as the *Euclidean distortion* of X. A fundamental result of Bourgain [4] states that $c_2(X) = O(\log n)$ for every n-point metric space (X, d).

For a general metric space (X, d) with no additional a-priori structure, there is a dearth of genuinely "interesting" constructions of Lipschitz mappings on X. One significant exception to this rule is provided by the distance functions $x \mapsto d(x, A)$ for some $\emptyset \neq A \subset X$. Of course, we can generate more examples by constructing Lipschitz functions to any finite dimensional normed space, the coordinates of which are distance functions. Observe that a mapping $f : X \to \ell_\infty$ is L-Lipschitz if and only if each of its coordinates is L-Lipschitz. These facts were put to good use in the classical observation of Fréchet that every metric space isometrically embeds into some $\ell_\infty(\Gamma)$ (see [3, 10]). Fréchet's embedding only uses distance functions for singleton A's, but more sophisticated refinements of this basic idea have appeared over the years. In what follows we call an embedding $f : X \to \mathbb{R}^{2^V \setminus \{\emptyset\}}$ a *Fréchet embedding* if for every $A \subset X$ there is $\alpha_A \in \mathbb{R}$ such that $f(x) = (\alpha_A d(x, A))_{A \in 2^X \setminus \{\emptyset\}}$.

Bourgain's embedding of finite metric spaces in ℓ_2 [4] is an instance of a Fréchet embedding in which the coefficients $\alpha_A \in [0,1]$ depends only on the cardinality of A. Bourgain's probabilistic method of producing a good Fréchet embedding has subsequently found many applications [6, 7, 9, 8, 5, 12].

As stated in the abstract, the present note is motivated by our recent Ramsey-type result:

Theorem 1 ([2]). *There exists an absolute constant $C > 0$ such that for every $\alpha > 2$, any finite metric M contains a subset $N \subset M$ for which $|N| \geq |M|^{1-C\frac{\log \alpha}{\alpha}}$ and $c_2(N) \leq \alpha$.*

The embedding used in Theorem 1 is not a Fréchet embedding. In view of the past success of Fréchet embeddings, and in particular Bourgain's embedding which gives the asymptotically best possible bound for embedding the whole metric space, it is natural to ask whether this (by now standard) method is applicable to Ramsey-type problems. This note provides a negative answer to this question. We find a certain range of the parameters, for which Fréchet embeddings fail to achieve tight bounds for Ramsey type questions:

Theorem 2. *For every $1/2 < \delta \leq 1$ there is a constant $C(\delta) > 0$ such that for infinitely many integers n there are n-point metric spaces M_n, such that for any Fréchet embedding $f : M_n \to \ell_p$, and any subset V of M_n of size at least n^δ, $dist(f|_V) \geq (C(\delta) \cdot \log n)^{1/p}$.*

We end the paper with a short discussion in which we comment on the embedding used in Theorem 1, showing that it can be viewed as a different natural generalization of Bourgain's embedding.

2 The Construction

Theorem 2 is proved by exhibiting an explicit example of an unbounded family of metric spaces for which every Fréchet embedding fails to yield the appropriate Ramsey type result. The example, denoted $Z = Z_{k,\varepsilon,h,m}$, is an amalgamation of two types of metric spaces:

- The elements of the first metric space, called $X = Ł_{h,k}$, are the leaves of a complete binary tree $\mathcal{T}_h$ of height h. The metric on the leaves is defined by

$$d_X(x_1, c_2) = k^{-l(\text{lca}(x_1,x_2))}.$$

  Here $\text{lca}(x_1, x_2)$ denotes the least common ancestor of x_1 and x_2 in $\mathcal{T}_h$ and $l(u)$ is the depth of the vertex u in T. The parameter $k > 1$ will be be fixed later. A crucial property of X is that for any $x, y, z \in X$, if $d_X(x,y) < d_X(x,z)$, then

$$d_X(x,z) = d_X(y,z) \geq kd_X(x,y) \tag{1}$$

  (Such a metric space is called a k-HST [1]).

- The second metric space, $Y = Y_{m,\varepsilon}$, is the one-dimensional metric on the points

$$Y = \left\{ y_0 = \varepsilon, y_1 = \varepsilon \left(1 + \tfrac{1}{4}\right), y_2 = \varepsilon \left(1 + \tfrac{1}{4} + \tfrac{1}{16}\right), \ldots, y_{m-1} = \varepsilon \left(\sum_{i=0}^{m-1} 4^{-i} \right) \right\}.$$
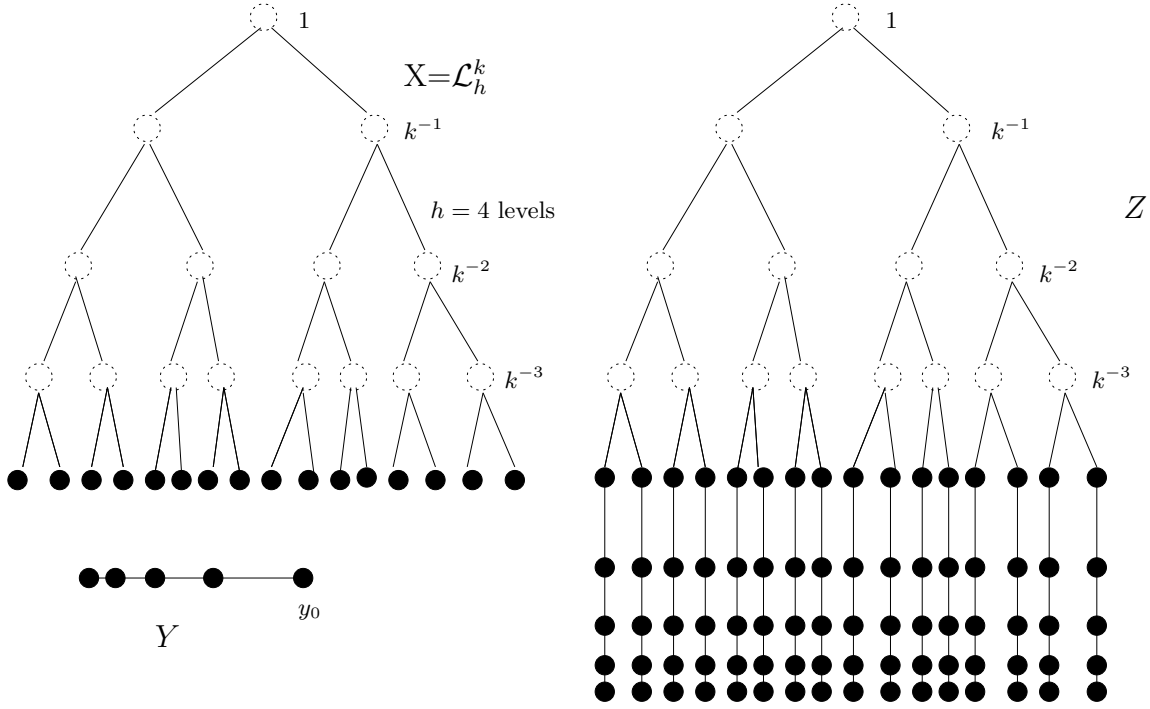
Figure 1: The metric space Z serving as the example in Theorem 2.

The points set of Z is $X \times Y$, hence its size is $2^h m$. The distance is defined by

$$d_Z((x,y),(x',y')) = \begin{cases} d_Y(y,y') & x = x' \\ d_Y(y,y_0) + d_X(x,x') + d_Y(y_0,y') & x \neq x'. \end{cases}$$

A schematic description of Z is given in Figure 1.

Here is a sketch of our proof. A preliminary step (Lemma 1) shows that all sufficiently large (of size $\geq n^{\frac{1}{2}+\epsilon}$) subsets $V \subseteq Z$ have a structure that is similar to the whole of Z. Namely, there is a large subset $U \subseteq V$ that spans a complete binary subtree (of a slight modification) of $\mathcal{T}$. Moreover, in each copy of Y that U meets, it contains at least two elements.

From this point on, we may assume, then, that we are dealing with a space that is similar to the original Z with two modifications: The depth has shrunk to one half its original value, (since we moved from n to $n^{\frac{1}{2}+\epsilon}$ points) and attached to each leaf is a *tail* i.e., a two-point metric space. On the other hand, we are now considering the whole of this (sub)space, and not subsets thereof. (We actually cannot ignore the points in $Z \setminus U$, which may participate in the subsets to which Fréchet embedding assigns nonzero weights. However, this is only a technicality and can be ignored at this level of discussion.)

Recall that we are dealing with Fréchet embeddings f that are defined by assigning weights α_A to subsets A. Consider a tail $\{y_1, y_2\}$, and suppose that A is disjoint from the copy of Y to which y_1, y_2 belong. Then $|d(y_1, A) - d(y_2, A)| = d(y_1, y_2)$, entailing a lower bound on $\mathrm{Lip}(f)$. In other words, if large weights α_A are allotted to subsets A that miss many tails, then $\mathrm{Lip}(f)$ is big. This is made precise in Lemma 2, which gives a lower bound on $\mathrm{Lip}(f)$ in terms of the numbers α_A and ζ_A - the fraction of tails that A misses.

It follows that in order to keep $\mathrm{Lip}(f)$ small, we ought to put significant weights α_A on subsets with large ζ_A i.e., those that meet many tails. This, however, tends to increase

Lip(f^{-1}), as we now explain. Let z_1 and z_2 belong to two distinct tails. In order for $|d(z_1, A) - d(z_2, A)|$ to be nearly equal to $d(z_1, z_2)$, the set A must meet exactly one of the two tails containing the z_i (in which case we say that A *splits* the least common ancestor of z_1, z_2). It is not hard to see (Equation 4) that if ζ_A is large, then A can split only a few vertices, and these necessarily reside far from the root. Note also that if A fails to split $\mathrm{lca}(z_1, z_2)$, then $|d(z_1, A) - d(z_2, A)| \leq \frac{1}{k} d(z_1, z_2)$, which is significant, since k is large. The precise argument is made in Lemma 4.

Omitting some additional technicalities (Lemma 3), these two considerations can be traded off against each other to yield the desired result.

Some notations and definitions concerning trees:

1. Let $\mathcal{T}$ be a tree. We denote its root by $r(\mathcal{T})$ and its set of leaves by $\mathrm{leaves}(\mathcal{T})$. For a subset $A \subset \mathrm{leaves}(\mathcal{T})$ let $\mathcal{T}(A)$ be the subtree spanned by A, i.e. the union of all the simple paths between elements of A and $\mathrm{lca}(A)$. The subtree $\mathcal{T}_u$ rooted at u consists of the union of all the monotone paths in $\mathcal{T}$ between u and its descendants. We recall that $l_{\mathcal{T}}(u)$ is the depth of u, i.e. its distance from $r(\mathcal{T})$.

2. Let $\mathcal{T} = (V, E)$ be a tree and suppose that $\mathcal{T}$ contains a vertex $u \neq r(\mathcal{T})$ with exactly two neighbors v and w. We *simplify* $\mathcal{T}$ by removing the vertex u and adding the edge (v, w). Another type of simplification step takes place, if $r(\mathcal{T})$ has a single neighbor v. In this case, we remove $r(\mathcal{T})$ and declare v as the root. The skeleton $\mathrm{skel}(\mathcal{T})$ is the (uniquely defined) tree obtained from $\mathcal{T}$ by carrying out all possible simplifications.

3. For a subset $U \subset Z$, we denote $U_X = \{x \in X; \ (\{x\} \times Y) \cap U \neq \emptyset\}$.

Lemma 1. *Let $h, m \geq 2$ be integers, and $Z = Z_{k,\varepsilon,h,m}$. Let $V \subset Z$ be such that $|V| > 2m + 2^h$. Denote*

$$\rho = \left(\frac{|V| - 2^h}{m} \right)^{1/h} - 1.$$

Then there is a subset $U \subset V$ such that for every $x \in U_X$, $|(\{x\} \times Y) \cap U| \geq 2$ and $\mathrm{skel}(\mathcal{T}_h(U_X))$ is a complete binary tree of depth at least $\frac{\rho}{6 \log(2/\rho)} h$.

Proof. Denote $S = \{x \in X; \ |V \cap (\{x\} \times Y)| \geq 2\}$. Since $|S| m + 2^h \geq |V|$, we have that

$$|S| \geq \frac{|V| - 2^h}{m} = (1 + \rho)^h > 2.$$

Set $0 < \eta = \frac{\rho}{2 + \rho} < 0.5$. We will first construct inductively a subtree $\mathcal{T}'$ of $\mathcal{T}$, as follows. Let r be the root of $\mathcal{T}$ and u_1, u_2 its two children. Recursively, assume that we have constructed $\mathcal{T}'_{u_1}, \mathcal{T}'_{u_2}$. $\mathcal{T}'$ is constructed according to the following rule: if $\min\{|\mathrm{leaves}(\mathcal{T}_{u_1}) \cap S|, |\mathrm{leaves}(\mathcal{T}_{u_2}) \cap S|\} \geq \eta |S|$ then $\mathcal{T}'$ is obtained by attaching $\mathcal{T}'_{u_1}$ and $\mathcal{T}'_{u_2}$ to r. Otherwise, assume without loss of generality that $|\mathrm{leaves}(\mathcal{T}_{u_1}) \cap S| > (1 - \eta)|S|$. In this case $\mathcal{T}'$ is obtained by attaching only $\mathcal{T}_{u_1}$ to r. Note that in the previous step of this procedure, $\mathcal{T}'_{u_i}$ were constructed with respect to sets $\mathrm{leaves}(\mathcal{T}_{u_i}) \cap S$.

Let u be a leaf of $\mathrm{skel}(\mathcal{T}')$ and denote its distance from the root of $\mathrm{skel}(\mathcal{T}')$ by λ. Consider the path in $\mathcal{T}$ between r (the root of $\mathcal{T}$) and u. Clearly λ is precisely the number of times along this path that the first option was taken in the above construction. We prove by induction on h, the height of $\mathcal{T}$, that $|S| \eta^\lambda (1 - \eta)^{h - \lambda} \leq 1$. The case $h = 0$ is trivial, so assume that $h > 0$.

4

Let r, u_1, u_2 be as above, and assume that u_1 belongs to the path connecting r and u. In the first case of the above construction, we get by the induction hypothesis that:

$$1 \geq |\text{leaves}(\mathcal{T}_{u_1}) \cap S| \eta^{\lambda-1}(1-\eta)^{h-1-(\lambda-1)} \geq |S|\eta^\lambda(1-\eta)^{h-\lambda}.$$

In the second case of the above construction we get

$$1 \geq |\text{leaves}(\mathcal{T}_{u_1}) \cap S| \eta^\lambda(1-\eta)^{h-1-\lambda} \geq |S|\eta^\lambda(1-\eta)^{h-\lambda}.$$

Let $\hat{h}$ be the minimal distance between a leaf of $\text{skel}(\mathcal{T}')$ and its root. From the discussion above (and noting that $\rho < 1$) it follows that:

$$\hat{h} \geq \frac{\log|S| + h\log(1-\eta)}{\log\left(\frac{1-\eta}{\eta}\right)} \geq \frac{h\log(1+\rho) + h\log(\frac{2}{2+\rho})}{\log\left(\frac{2}{\rho}\right)} \geq \frac{\rho}{6\log(2/\rho)}h.$$

Consider all the vertices of depth exactly $\hat{h}$ in $\text{skel}(\mathcal{T}')$. The tree $\text{skel}(\mathcal{T}')$ truncated at these points forms a complete binary tree of height $\hat{h}$. For each such vertex u we choose some $s_u \in S \cap \mathcal{T}'_u$. The set $S' = \{s_u; \ u \text{ has depth } \hat{h} \text{ in } \text{skel}(\mathcal{T}')\}$ is easily seen to have the required properties. $\qquad\square$

Our next two lemmas provide a lower bound on the expansion of any Fréchet embedding $Z \hookrightarrow \ell_p$, using (almost) only the structure of Y.

Let $f : Z \to \ell_p^{2^Z \setminus \{\emptyset\}}$, $1 \leq p < \infty$, be a Fréchet embedding given by $f(x)_A = \alpha_A d(x, A)$, let $V \subseteq Z$ be a large set as in Theorem 2 and Let $U \subset V$ be a subset as in Lemma 1. For every $\emptyset \neq A \subset Z$ write:

$$\zeta_A = \frac{|\{x \in U_X; \ (\{x\} \times Y) \cap A = \emptyset\}|}{|U_X|}.$$

Namely, ζ_A is the fraction of of the tails in U that A misses. Define

$$\beta(f, U) = \left(\sum_{\emptyset \neq A \subset Z} |\alpha_A|^p \zeta_A\right)^{1/p}. \tag{2}$$

Informally, $\beta(f)$ is a weighted average of the fraction of tails $\{x\} \times Y$, $x \in U$, that are missed by a subset of Z. If a subset A is disjoint from a tail $\{x\} \times Y$, that contains the points $(x, u), (x, v)$, then $|d_Z((x, u), A) - d_Z((x, v), A)| = d_Y(u, v)$. Consequently, we can expect to obtain lower bound on the Lipschitz constant of f in terms of β. Indeed:

Lemma 2. Let $U \subset Z$ be a subset as in Lemma 1. Then $\text{Lip}(f|_U) \geq \beta(f, U)$.

Proof. We may assume without loss of generality that $\alpha_A \geq 0$ for every $\emptyset \neq A \subset Z$. Set $\beta = \beta(f, U)$. Note that

$$\frac{1}{|U_X|} \sum_{x \in U_X} \left(\sum_{A: \ A \cap Z_x = \emptyset} \alpha_A^p\right) = \sum_{\emptyset \neq A \subset Z} \alpha_A^p \frac{|\{x \in U_X; \ A \cap (\{x\} \times Y) = \emptyset\}|}{|U_X|} = \beta^p.$$

Hence there is some $x \in U_X$ such that $\sum_{A: \ A \cap Z_x = \emptyset} \alpha_A^p \geq \beta^p$. As mentioned, if $\emptyset \neq A \subset Z$ is disjoint from $\{x\} \times Y$, then $d_Z((x, y), A) = d_Y(y, y_0) + d_Z((x, y_0), A)$ for every $(x, y) \in Z$.

Recall that for $x \in U_X$, there are $y \neq y'$ such that $(x, y), (x, y') \in (\{x\} \times Y) \cap U$. Therefore

$$\|f(x, y) - f(x, y')\|_p^p = \sum_{A \subset Z} \alpha_A^p |d_Z((x, y), A) - d_Z((x, y'), A)|^p$$

$$\geq \sum_{\substack{A \subset Z \\ A \cap (\{x\} \times Y) = \emptyset}} \alpha_A^p |d_Y(y_0, y) - d_Y(y_0, y')|^p \geq \beta^p d_Z((x, y), (x, y'))^p,$$

so that $\text{Lip}(f|_U) \geq \beta$. $\qquad \square$

Similar to the definition of $\beta(f, U)$, we also define

$$\gamma(f, U) = \left(\sum_{\emptyset \neq A \not\supset U} \alpha_A^p \right)^{1/p}. \qquad (3)$$

In words, $\gamma(f, U)$ is the weighted average of weights assigned to subsets that do not contain U. The metric of Y, was purposely chosen so that no two distances are equal. Consequently, a subset $A \not\supset U$ makes a nonzero contribution to $\text{Lip}(f|_U)$. Indeed:

Lemma 3. *Let $U \subseteq Z$ be as in Lemma 2. Assume that $k \geq 4$, and $k^{-h} \geq \varepsilon$. Then $\text{Lip}(f|_U) \geq 4^{-m} 2^{-\frac{|X| \cdot |Y|}{p}} \gamma(f, U)$.*

Proof. We may assume again that $\alpha_A \geq 0$ for every $\emptyset \neq A \subset Z$. Let A_0 be such that $\alpha_{A_0} = \max\{\alpha_A; \emptyset \neq A \not\supset U\}$. Clearly:

$$\alpha_{A_0}^p \geq 2^{-|X| \cdot |Y|} \sum_{\emptyset \neq A \not\supset V} \alpha_A^p.$$

Since $A_0 \not\supset U$, there is some $(x, u) \in U \setminus A_0$. By assumption, U contains some (x, v) with $v \neq u$. It is enough to show that $|d_Z((x, u), A_0) - d_Z((x, v), A_0)| \geq 4^{-m} \cdot \text{diam}(Y)$, since then:

$$\|f(x, u) - f(x, v)\|_p \geq \alpha_{A_0} |d_Z((x, u), A_0) - d_Z((x, v), A_0)| \geq 4^{-m} \alpha_{A_0} \text{diam}(Y) \geq 4^{-m} \alpha_{A_0} d_Y(u, v),$$

and the result will follow from the lower estimate for α_{A_0}.

To verify that $|d_Z((x, u), A_0) - d_Z((x, v), A_0)| \geq 4^{-m} \cdot \text{diam}(Y)$, distinguish between the possible points where the distances $d_Z((x, u), A_0)$ and $d_Z((x, v), A_0)$ are attained. Let $(x', a) \in A_0$, $(x'', b) \in A_0$ satisfy $d_Z((x, u), A_0) = d_Z((x, u), (x', a))$, $d_Z((x, v), A_0) = d_Z((x, v), (x'', b))$. Observe that since $k^{-h} > \varepsilon/3 > \text{diam}(Y)$, if $x \in \{x', x''\}$ then $x' = x'' = x$, in which case $d_Z((x, u), A_0) = d_Y(u, a)$, $d_Z((x, v), A_0) = d_Y(v, b)$, and we conclude since the definition of the metric on Y implies that $|d_Y(u, a) - d_Y(u, b)| \geq 4^{-m+1} \varepsilon > 4^{-m} \text{diam}(Y)$. Therefore, assuming that $x \notin \{x', x''\}$, $d_Z((x, u), A_0) = d_Y(u, y_0) + d_X(x, x') + d_Y(y_0, a)$ and $d_Z((x, v), A_0) = d_Y(v, y_0) + d_X(x, x'') + d_Y(y_0, b)$. If $d_X(x, x') = d_X(x, x'')$, we use the fact that if $a, b, u, v \in Y$ but $\{a, b\} \neq \{u, v\}$ then $|[d_Y(u, y_0) + d_Y(y_0, a)] - [d_Y(v, y_0) + d_Y(y_0, b)]| \geq 4^{-m} \text{diam}(Y)$. On the other hand, if $d_X(x, x') \neq d_X(x, x'')$ then $|d_X(x, x') - d_X(x, x'')| \geq k^{-h} \geq \varepsilon$, and so

$$\begin{aligned}
|d_Z((x, u), A_0) - d_Z((x, v), A_0)| &\geq |d_X(x, x') - d_X(x, x'')| - \\
&\quad |d_Y(u, y_0) - d_Y(v, y_0)| - |d_Y(a, y_0) - d_Y(b, y_0)| \\
&\geq \varepsilon - 2\text{diam}(Y) \geq \text{diam}(Y).
\end{aligned}$$

$\qquad \square$

Our next lemma uses the special structure of $X = \mathcal{L}_{h,k}$ to bound from below the inverse of the contraction, $\text{Lip}((f|_U)^{-1})$, in terms of $\beta(f, U)$ and $\gamma(f, U)$. In what follows we always use the convention that $\text{Lip}(f^{-1}) = \infty$ if f is non injective.

Lemma 4. *Let $U \subset Z$ be as in Lemma 1 . Then for every Fréchet embedding $f : Z \to \ell_p^{2^Z \setminus \{\emptyset\}}$,*

$$\text{Lip}\left((f|_U)^{-1}\right) \geq \left[\frac{2\beta(f, U)^p}{\hat{h}} + \left(\frac{2 + 2k^h \operatorname{diam}(Y)}{k}\right)^p \gamma(f, U)^p\right]^{-1/p}.$$

Proof. We use a shorthand notation $\mathcal{T} = \mathcal{T}_h$, $\gamma = \gamma(f, U)$ and $\beta = \beta(f, U)$. We assume as usual that $\alpha_A \geq 0$ for every $\emptyset \neq A \subset Z$. Denote by $\mathcal{T}' = \mathcal{T}_h(U_X)$ the subtree generated by U_X. By our assumption $\text{skel}(\mathcal{T}')$ is a complete binary tree of depth $\hat{h}$. Let $u; u_1, u_2$ be a vertex in $\text{skel}(\mathcal{T}')$ and its two children in $\mathcal{T}$. We say that a subset $A \subset Z$ *splits* u if A intersects exactly one of the sets $\text{leaves}(\mathcal{T}_{u_1}) \times Y, \text{leaves}(\mathcal{T}_{u_2}) \times Y$.

For $u \in \text{skel}(\mathcal{T}')$, denote by $l(u) = l_{\text{skel}(\mathcal{T}')}(u)$ the depth of u in $\text{skel}(\mathcal{T}')$. We first claim that for $\emptyset \neq A \subset Z$,

$$\sum_{\substack{u \in \text{skel}(\mathcal{T}') \\ A \text{ splits } u}} 2^{-l(u)} \leq 2\zeta_A. \tag{4}$$

The proof is by induction on $\hat{h}$, the height of $\text{skel}(\mathcal{T}')$. Let r be the root of $\text{skel}(\mathcal{T}')$. Denote by u_1, u_2 the children of r in $\mathcal{T}$, and by v_1, v_2 the children of r in $\text{skel}(\mathcal{T}')$. Set $\zeta = \zeta_A$ and define ζ_1, ζ_2 by:

$$\zeta_i = \frac{|\{x \in \text{leaves}(\mathcal{T}'_{v_i}); (\{x\} \times Y) \cap A = \emptyset\}|}{|\text{leaves}(\mathcal{T}'_{v_i})|}, \quad i = 1, 2.$$

Note that since $\text{skel}(\mathcal{T}')$ is a complete binary tree $\zeta = (\zeta_1 + \zeta_2)/2$. If A does not split r then by induction:

$$\sum_{\substack{u \in \text{skel}(\mathcal{T}') \\ A \text{ splits } u}} 2^{-l(u)} = \frac{1}{2}\left(\sum_{\substack{u \in \text{skel}(\mathcal{T}'_{v_1}) \\ A \text{ splits } u}} 2^{-(l(u)-1)} + \sum_{\substack{u \in \text{skel}(\mathcal{T}'_{v_2}) \\ A \text{ splits } u}} 2^{-(l(u)-1)}\right)$$

$$\leq \frac{1}{2}(2\zeta_1 + 2\zeta_2) = 2\zeta.$$

On the other hand, if A splits r, it does not intersect one of $\text{leaves}(\mathcal{T}_{u_1}) \times Y$, $\text{leaves}(\mathcal{T}_{u_2}) \times Y$, say it does not intersect $\text{leaves}(\mathcal{T}_{u_2}) \times Y$. In this case $\zeta_2 = 1$ so that by the induction hypothesis:

$$\sum_{\substack{u \in \text{skel}(\mathcal{T}') \\ A \text{ splits } u}} 2^{-l(u)} = 1 + \frac{1}{2}\left(\sum_{\substack{u \in \text{skel}(\mathcal{T}'_{v_1}) \\ A \text{ splits } u}} 2^{-(l(u)-1)}\right) \leq 1 + \frac{1}{2}2\zeta_1 = \zeta_2 + \zeta_1 = 2\zeta.$$

This finishes the proof of (4). Now, by the definition of $\beta = \beta(f, U)$,

$$\sum_{l=0}^{\hat{h}-1} \sum_{\substack{u \in \text{skel}(\mathcal{T}') \\ u \text{ in depth } l \\ \text{in skel}(\mathcal{T}')}} 2^{-l} \sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ splits } u}} \alpha_A^p = \sum_{\emptyset \neq A \subset Z} \alpha_A^p \sum_{\substack{u \in \text{skel}(\mathcal{T}') \\ A \text{ splits } u}} 2^{-l(u)} \leq \sum_{\emptyset \neq A \subset Z} \alpha_A^p 2\zeta_A = 2\beta^p.$$

It follows that there exists some $l \in \{0, \ldots, \hat{h} - 1\}$ such that

$$2^{-l} \sum_{\substack{u \in \mathrm{skel}(\mathcal{T}') \\ u \text{ in depth } l \\ \text{in } \mathrm{skel}(\mathcal{T}')}} \left(\sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ splits } u}} \alpha_A^p \right) \leq \frac{2\beta^p}{\hat{h}}.$$

So there exists a vertex $u \in \mathrm{skel}(\mathcal{T}')$ (in depth l in $\mathrm{skel}(\mathcal{T}')$) such that

$$\sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ splits } u}} \alpha_A^p \leq \frac{2\beta^p}{\hat{h}}.$$

Denote by u_1, u_2 the two children of u **in** $\mathcal{T}$. Since u is a vertex of $\mathrm{skel}(\mathcal{T}')$ there are $x_1 \in \mathrm{leaves}(\mathcal{T}_{u_1}) \cap U_X$ and $x_2 \in \mathrm{leaves}(\mathcal{T}_{u_2}) \cap U_X$ (in particular $\mathrm{lca}(x_1, x_2) = u$). Fix $(x_1, y_1) \in U$ and $(x_2, y_2) \in U$. The observation here is that if a subset $A \subset Z$ does not split u then one of the following two cases happens:

1. $A \cap (\mathcal{T}_u \times Y) = \emptyset$. In this case, $d_Z((x_1, y_1), A) = d_Z((x_2, y_2), A)$, and therefore

$$|d_Z((x_1, y_1), A) - d_Z((x_2, y_2), A)| = 0.$$

2. For $i = 1, 2$ there is $x_i' \in \mathrm{leaves}(\mathcal{T}_{u_i})$ and $y_i', y_i'' \in Y$ such that $d_Z((x_i, y_i), A) = d_Y(y_i, y_i') + d_X(x_i, x_i') + d_Y(y_0, y_i'')$. Hence,

$$\begin{aligned}
|d_Z((x_1, y_1), A) - d_Z((x_2, y_2), A)| &\leq d_X(x_1, x_1') + d_X(x_2, x_2') + 2\,\mathrm{diam}(Y) \\
&\leq \frac{2}{k} d_X(x_1, x_2) + 2\,\mathrm{diam}(Y) \\
&\leq \frac{2 + 2k^h\,\mathrm{diam}(Y)}{k} d_Z((x_1, y_1), (x_2, y_2)).
\end{aligned}$$

Therefore

$$\begin{aligned}
\|f((x_1, y_1)) - &f((x_2, y_2))\|_p^p = \\
&= \sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ splits } u}} \alpha_A^p |d_Z((x_1, y_1), A) - d_Z((x_2, y_2), A)|^p \\
&\quad + \sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ doesn't splits } u}} \alpha_A^p |d_Z((x_1, y_1), A) - d_Z((x_2, y_2), A)|^p \\
&\leq [d_Z((x_1, y_1), (x_2, y_2))]^p \sum_{\substack{\emptyset \neq A \subset Z \\ A \text{ splits } u}} \alpha_A^p \\
&\quad + \left(\frac{2 + 2k^h\,\mathrm{diam}(Y)}{k} d_Z((x_1, y_1), (x_2, y_2)) \right)^p \sum_{\emptyset \neq A \not\supset U} \alpha_A^p \\
&\leq \left[\frac{2\beta^p}{\hat{h}} + \left(\frac{2 + 2k^h\,\mathrm{diam}(Y)}{k} \right)^p \gamma^p \right] [d_Z((x_1, y_1), (x_2, y_2))]^p.
\end{aligned}$$

$\square$

8

Proof of Theorem 2. Let $h, m \geq 2$ where $\epsilon = k^{-h}/2$ Define $n = m2^h$, so that $|Z| = n$.

Let $f : Z \to \ell_p^{2^Z \setminus \{\emptyset\}}$ be a Fréchet embedding. Fix $0.5 < \delta < 1$. We can always choose h and m such that $n^\delta = \Omega(2^h + mn^{2\delta-1})$ (2^h of order n^δ and m of order $n^{1-\delta}$). Fix such h and m, and let $V \subset Z$ be such that $|V| \geq n^\delta$. By Lemma 1 there is a subset $U \subset V$ which satisfies the conditions of Lemma 4 with $\hat{h} \geq C(\delta) \log n$. Set $\beta = \beta(f, U)$, $\gamma = \gamma(f, U)$. It follows from Lemmas 2, 3, 4 that:

$$\text{dist}(f|_V) \geq \text{dist}(f|_U) \geq c \frac{\max\{\beta, 2^{-n/p}4^{-m}\gamma\}}{\max\left\{\frac{\beta}{(C(\delta)\log n)^{1/p}}, \frac{\gamma}{k}\right\}},$$

for some universal constant $c > 0$.

If $\frac{\beta}{(C(\delta)\log n)^{1/p}} \geq \frac{\gamma}{k}$ then we deduce that $\text{dist}(f|_V) \geq (C(\delta) \log n)^{1/p}$, as required. Otherwise we get the lower bound $\text{dist}(f|_V) \geq ck2^{-n/p}4^{-m}$. Recall that we are still free to choose k, so that the required result also follows from this case provided that k is large enough. $\square$

3   Discussion

The goal of Theorem 2 has been to provide an example for which Fréchet embeddings fail to achieve the best possible bounds. This is done for the problem of embedding subsets of size at least n^δ into ℓ_p, where $\delta > 1/2$. We suspect that the same holds also for $\delta \leq 1/2$. However, the example presented here does not immediately apply to this case.

Finally we comment on a concept due to Matoušek and Rabinovich [11] which is a different natural generalization of Bourgain's embedding. Given a finite metric space (V, ρ), we say that a one-dimensional metric σ on V is *dominated* by ρ if $\rho(x, y) \geq \sigma(x, y)$ for every $x, y \in V$. The polytope $\ell_1^{\text{dom}}(\rho)$ is the convex hull of all one-dimensional metrics dominated by ρ. It is natural to ask whether embeddings into ℓ_1^{dom} can be used for Ramsey-type problems. We observe that this is indeed the case. The following theorem is a consequence of the main Ramsey-type theorem of [2]:

Theorem 3. *For every finite metric M, and every $\alpha > 2$, there exists a subset $N \subset M$ of cardinality $\geq |M|^{1-C\frac{\log \alpha}{\alpha}}$ that is $O(\alpha)$-equivalent to some metric in $\ell_1^{\text{dom}}(N)$. Here $C > 0$ is an absolute constant.*

The proof of Theorem 3 is in two steps: (i) We recall that Theorem 1 follows from a Ramsey-type theorem of [2] where the target metric space for the embedding is an *ultrametric* (which is isometrically embeddable in ℓ_2). (ii) Every ultrametric ρ is $O(1)$-equivalent to some metric in $\ell_1^{\text{dom}}(\rho)$. This is a simple fact, a proof of which, is sketched below.

Proposition 5. *Every ultrametric ρ is 16-equivalent to some metric in $\ell_1^{\text{dom}}(\rho)$.*

Sketch of a proof. As shown in [2], ρ has a 4-embedding in some 4-HST. So, it suffices to prove that ρ is 4 equivalent to some metric in $\ell_1^{\text{dom}}(\rho)$ for the case where ρ itself is a 4-HST metric. Let us recall: A k-HST is defined by a tree T in which every internal vertex v is assigned a weight $\Delta(v) > 0$. If v is a child of u then $\Delta(u) \geq k\Delta(v)$. The metric is defined of T's leaves via $d(x, y) = \Delta(\text{lca}(x, y))$. The metric in $\ell_1^{\text{dom}}(\rho)$ that is 4-equivalent to ρ is constructed through a probabilistic argument. Associate with every edge $e = uv$ a weight $\epsilon_e \in \{-\frac{3}{8}, \frac{3}{8}\}$. This is done uniformly and independently over all edges. Associated with this is a mapping $\varphi = \varphi_\epsilon$ of T's leaves to the real line, $\varphi(x) = \sum \Delta(u)\epsilon_{uv}$. This sum extends over all edges uv in the directed path from T's root to x. It is easy to verify that the one dimensional metric induced on the leaves is dominated by ρ. This is now averaged over all possible choices of the ϵ_e's. $\square$

References

[1] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science*, pages 184–193, 1996.

[2] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On Metric Ramsey-type Phenomena. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, 2003.

[3] Y. Benyamini and J. Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[4] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[5] U. Feige. Approximating the bandwidth via volume respecting embeddings. *J. Comput. System Sci.*, 60(3):510–539, 2000.

[6] W. B. Johnson, J. Lindenstrauss, and G. Schechtman. On Lipschitz embedding of finite metric spaces in low-dimensional normed spaces. In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 177–184. Springer, Berlin, 1987.

[7] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[8] J. Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.*, 93:333–344, 1996.

[9] J. Matoušek. On embedding expanders into l_p spaces. *Israel J. Math.*, 102:189–197, 1997.

[10] J. Matoušek. *Lectures on discrete geometry*. Springer-Verlag, New York, 2002.

[11] J. Matoušek and Yuri Rabinovich. On dominated l_1 metrics. *Israel J. Math.*, 123:285–301, 2001.

[12] M. Thorup and U. Zwick. Approximate distance oracles. In *33rd Annual ACM Symposium on Theory of Computing*, pages 183–192, 2001.

Yair Bartal, Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel.
yair@cs.huji.ac.il

Nathan Linial, Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel.
nati@cs.huji.ac.il

Manor Mendel, Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel.
mendelma@cs.huji.ac.il

Assaf Naor, Theory Group, Microsoft Research, One Microsoft Way 113/2131, Redmond WA 98052-6399, USA.
anaor@microsoft.com